

MICHAEL C. ORMSBY
United States Attorney
Timothy M. Durkin
Assistant United States Attorney
Eastern District of Washington
Post Office Box 1494
Spokane, WA 99210-1494
Telephone: (509) 353-2767
Victor Boutros, Trial Attorney
U.S. Department of Justice
Civil Rights Division – Criminal Section
950 Pennsylvania Ave., NW
Washington, D.C. 20530
Tel. (202) 514-3204

SR. JUDGE FRED VAN SICKLE

UNITED STATES DISTRICT COURT
EASTERN DISTRICT OF WASHINGTON

I, LISA FOURNIER, declare, under penalty of perjury under the laws of the United States, that the following is true and correct to the best of my knowledge:

1. I am over the age of 18 and competent to testify herein.
2. This declaration is based upon my own personal knowledge and academic training.

Declaration of Dr. Fournier
Page 1

1 This declaration was prepared by **Dr. Lisa Fournier, PhD**. I am currently an
2 associate professor and the director of experimental training in the psychology
3 department at Washington State University. I received my PhD in Experimental
4 Psychology at the University of Illinois at Urbana-Champaign in 1993 and
5 completed a two-year post doc in the cognitive psychophysiology at the same
6 university. My research is focused primarily on attention, memory, perception
7 and action. The courses I currently teach include research methods (which covers
8 research design, hypothesis testing, measurement, internal and external
9 validity/reliability, and ethics), history of psychology, and a graduate course in
10 cognitive psychology. I am currently an associate editor for two psychology
11 journals (peer-reviewed), a consultant for the USAF, and I have published 23
12 articles in peer-reviewed journals.

13 I was asked by John Drake from the United States Attorneys Office to peer review
14 the research provided in the documents below in which William Lewinski is an
15 author.

16 **The materials used for this declaration include the following, sent to me by John
17 Drake from the United States Attorneys Office:**

- 18 1. A copy of an article titled “The attention study: A study on the presence of
19 selective attention in firearms officers” by William Lewinski published in *Law
Enforcement Executive Forum* (2008).
- 20 2. A copy of a newsletter from “Force Science News” #176 and an online link to
21 the Discovery Channel which features “The Force Science Exhaustion Study”
22 conducted by William Lewinski and colleagues which attempts to examine
23 the relationship between memory recall and exhaustion.
- 24 3. A copy of a magazine article titled “Biomechanics of lethal force
25 encounters—officer movements” by W. Lewinski published in *The Police
Marksman* magazine (2002).
- 26 4. A copy of a magazine article titled “shooting drills for the police sniper” by
27 Lewinski & Hudson, published in *The Police Marksman* magazine (2003).
- 28 5. A copy of an article titled “An examination of police officer mental
29 chronometry: “I swear... I don’t know how I shot him in the back” by
30 Bumgarner, Lewinski, Hudson, and Sapp published in *The Scene: Journal of
the Association for Crime Reconstruction* (date unknown)

31
32 Declaration of Dr. Fournier
33 Page 2

6. A copy of a paper titled “Biomechanics of lethal force encounters: Officer’s and subject’s movement and speed (a compilation of three studies) by Lewinski and Wilkins presented at the *Academy of Criminal Justice Sciences Annual Conference* in Boston, MA (2003).
7. A copy of an article titled “Command types used in police encounters” by Schwarzkopf, Houlihan, Kolb, Lewinski, Buchanan, and Christenson published in *Force Science Forum* (a medium for delivery of research conducted and supported by the Force Science Research Center headed by William Lewinski).
8. A copy of William Lewinski’s dissertation published by University Microfilms International as partial fulfillment for a PhD in Police Psychology from Union Institute & University in Cincinnati (Which I believe is an adult online degree program).
9. A copy of an article titled “A survey of the research on human factors related to lethal force encounters: Implications for law enforcement training, tactics, and testimony” by Honig & Lewinski published in *Law Enforcement Executive Forum* (2008).

Other information sent to me by John Drake from the United States Attorneys Office includes:

A copy of the statement of Karl F. Thompson (RPT: 06-79962) based on an interview by Detective T. Ferguson, surveillance video clips taken from the "Zip Trip", and a copy of William Lewinski's Expert Report.

Declaration of Dr. Fournier
Page 3

1 Declaration

2 Based on my review of the research conducted by William Lewinski (1-8 and a
3 survey article 9) listed above, I have concluded that this research (with the
4 exception of 7 above) is not internally valid or reliable due to 1) flaws (confounds)
5 in research methodology, 2) the lack of statistical comparisons among conditions,
6 and/or 3) conclusions drawn based on manipulations that were not incorporated
7 in the research design. Surprisingly, most of the research described below ignores
8 basic concepts in research design, hypothesis testing, internal validity and
9 reliability which are concepts covered in an undergraduate research design
10 course. A brief summary of why each of the above studies is invalid/unreliable is
11 presented below. A more detailed summary is provided (in attachments) for
12 studies that had multiple validity/reliability problems, and that may be used as
13 evidence to support William Lewinski's expert testimony.

14 First, in the four studies (3, 4, 5, and 6 listed above) that investigate
15 biomechanics of lethal force encounters, shooting drills, movement and speed,
16 and mental chronometry, there are no statistical comparisons to determine
17 whether one type of motion/reaction time event was significantly, and hence
18 reliably, different from another. When comparing mean reaction times (or any
19 other measure based on means), one must take into account the variability
20 around the mean to determine whether one mean is significantly different than
21 another. Also one must show that the probability that this difference is due to
22 chance is very small (less than 5%). Because no such statistics were reported, one
23 cannot claim that the results and the conclusions drawn from the results in these
24 articles are valid or reliable. Unfortunately the author(s) of these articles make
25 several unsubstantiated claims about differences in movement times/trigger
26 reaction times in different event scenarios even though no such statistical
27 analyses were conducted. The importance of reporting statistics (accurate
statistics) to show that your findings are valid/reliable is basic knowledge learned
in an undergraduate research methods course. Unfortunately, similar problems
exist in most of the other research reviewed below.

28 Second, in the study (1 listed above) investigating the focus of attention in
29 firearm officers, there were several flaws (confounds) in this study that make the
30 study invalid and unreliable. Because several major flaws were found in the
31

32 Declaration of Dr. Fournier
33 Page 4

1 experimental design and because many of the conclusions that were drawn by
2 the author were not possible based on the experimental design and analyses, a
3 more detailed review of this article is attached (see Attachment 1). In brief: 1)
4 There were no statistical comparisons of the mean frequencies for incidences
5 recalled or for recall errors assigned to different “attentional processes”, and
6 hence one cannot reliably claim that officers were engaged in one particular
7 “attentional process” more frequently than another (e.g., external focused
8 attention vs. external broad attention). 2) There was no control condition or
9 control group in which attention (assessed by memory recall) was measured
10 during a non-conflict event to serve as a baseline to determine whether a conflict
11 event influenced/changed/debilitated the allocation of attention relative to a
12 non-conflict situation. Thus, any statements suggesting that attention (or
13 memory) was affected differently in the high conflict situation (or that the
14 stressful situation in this scenario affects attention in any way that is different
15 from any other scenario) is not valid based on this data. 3) The statement that
16 ***“officers were externally tunneled specifically on items/behavior that were
17 crucial for their performance”*** is not valid or reliable because memory recall data
18 was not coded or analyzed based on whether it was “crucial” or “not crucial” to
19 performance—all memory recall data was recorded regardless of this distinction.
20 4) There was no mention or statistical analysis showing that the three “graders”
21 had assigned each of the memory-recall events from each officer consistently to
22 the different “attentional process” categories. That is, inter-rater reliability was
23 not evaluated to ensure reliability of the measure. 5) It is important to mention
24 that while accurate memory recall for events suggests that this information was
25 previously attended, it does not mean that items which are not recalled from
26 memory were not attended. The author erroneously makes the claim that what is
27 recalled or not recalled (or not correctly recalled) can be used as a basis of what
was and was not attended. Such a claim is not valid or reliable based on the
methods used in this study because memory recall was assessed which does not
directly assess the allocation of attention. Although there are other problems
concerning the sample size and unequal observations in each condition that also
challenge the validity of this study and reliability of the data, simply having one of
the violations described above is enough to show that the data and
interpretations drawn from these data are unreliable.

28
29 Third, the Force Science Exhaustion Study (2 listed above—Force Science
30 newsletter) failed to statistically analyze differences in the frequency of memories
31

32 Declaration of Dr. Fournier
33 Page 5

1 accurately reported by officers that were exerted (exerters) compared to officers
2 that were not exerted (control condition) during a “threat scenario”. Thus, the
3 authors’ claim that attention or memory is influenced by fatigue/exertion is not
4 valid or reliable. Also, the authors’ suggest that “the *more* exhausted officers
5 were, the less accurate their estimates [memories] tended to be”. Importantly,
6 there was no data or statistics to validate this claim. Because this study is
7 potentially relevant to the current case, a more detailed analysis of this study is
8 attached (see **Attachment 2**). Moreover, the officers ages varied greatly and
9 were not matched in terms of age or baseline heart rate before assigning them to
10 the different conditions (e.g., exerter/non-exerter). Because the samples
11 represented in the different conditions differed before the start of the
12 experiment, it is not possible to draw any conclusions between these conditions.
13 This major confound makes this data and any conclusions drawn from it invalid
14 and unreliable. Finally, because this study is only assessing memory, one cannot
15 make a statement about what was NOT attended. It is possible that information
16 was attended, but was forgotten. In other words, what is explicitly remembered
17 does indicate that this information was attended, but what is attended does not
18 determine what will be remembered. However, the authors claim that what was
19 “filtered out” or inhibited by attention can be determined by what was not
20 remembered. In summary, one can not reliably draw conclusions from this study
21 due to lack of reliability established through statistical analyses and due to some
22 of the claims made concerning attention that cannot be validated since memory
23 (not attention) was measured—these two cognitive processes are not identical.
24 Even if the results from this study had been reliable, it could not be reliably used
25 for the current case as memory was assessed immediately after the “threat”
26 scenario.
27

28 Fourth, I provided a peer review of William Lewinski’s published dissertation
29 (see **Attachment 3**). I believe that this dissertation demonstrates a lack of
30 knowledge William Lewinski has concerning research design and a lack of
31 understanding he has in terms of what can and cannot be concluded based on
32 one’s research design. The study presented in his dissertation violates many of
33 the design, biases, and validity/reliability issues covered in an undergraduate
34 research methods course. Not only would such a dissertation be rejected in my
35 department, it would not receive a passing grade in my undergraduate research
36 methods course.
37

38 Declaration of Dr. Fournier
39 Page 6

1 Fifth, my peer evaluation of the article discussing command types used in
2 police encounters is overall positive. I found no obvious threats to internal
3 validity or reliability. The authors presented appropriate data analyses,
4 conclusions based on the design, inter-rater reliability measures, and discussed
5 potential limitations of the design. This study was in stark contrast to the other
6 studies discussed above. (Note that in this study, Mr. Lewinski was not among
7 the first few authors, and hence likely did not contribute as much to this study
compared to the others).

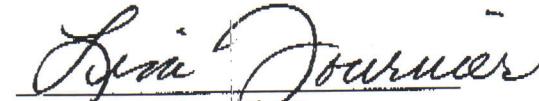
8 Sixth, the survey research on lethal force (listed as 9 above) had a few
9 unreliable statements because they were based on research (described above)
10 shown to be unreliable or on not-yet published research (and hence its reliability
11 could not be evaluated). For example, on page 130 data is cited from Lewinski
12 and Hudson (2003) and false interpretations based on these unanalyzed data are
13 incorporated in the text (see "lessons learned" on page. 130). Also, there is no
14 published data to support the strong claim that: failing to identify a weapon
15 when one exists results in an increase in seeing a weapon when none exists (page
16 130). Furthermore, there was no empirical data to support the statement, "One
17 type of memory can be used to stimulate another. For example, after a shooting,
18 a walk-through would stimulate recognition memory, which would then facilitate
19 recall memory and provide a more accurate and thorough report of the incident
by the officer." (p. 132). However, I found most of the summary of research
related to attention and memory to be accurate (based on my knowledge of the
literature) with the exception of the recommendations made by the authors in
terms of how to reduce memory errors—which was not supported by reliable,
empirical research. Because this was survey research, most of the information
was sound and based on empirical research published in peer-reviewed journals.

20 In summary, although the author(s) in the above studies (except for survey
21 article, labeled 9 above and the study on command types, labeled 7 above) often
22 describe results that we may believe are logical, there are no empirical bases
23 underlying them. Drawing scientific conclusions on logic alone, without empirical
24 data based on good research practices and appropriate analyses, is nothing more
25 than pseudoscience. It is my opinion that the data reported in these published
26 studies are not internally valid nor are the reliable. The research studies above
27 (except for 7) ignore basic concepts in research design, hypothesis testing,
internal validity and reliability which are concepts covered in an undergraduate

Declaration of Dr. Fournier
Page 7

1 research design course. As a result, I do not believe they can be applied reliably
2 to the current court case or any other case in the court of law.

3 DATED this 23^r day of October, 2011.

4
5 
6 Lisa Fournier, Ph.D.

Attachment 1 (1 of 2 pages)

Article titled "The Attention Study: A Study on the Presence of Selective Attention in Firearms Officers" by William Lewinski published in *Law Enforcement Executive Forum* (2008).

General analysis of this manuscript. The claims in this paper are not valid or reliable due to flaws (confounds) in methodology, lack of statistical comparisons among conditions, and conclusions claimed based on manipulations that were not incorporated in the research design.

The main **purpose** of this study was to determine the extent that selective attention and inattentional blindness occurs in well-trained, armed response officers in a conflict scenario involving a simulated armed encounter. The research also examines memory recall of an event using "different means of acquiring information" from these officers. **Method.** Groups of 2-3 fire-arm officers were exposed to a "conflict" scenario in which they had to disable a person pointing a shotgun at them. After disabling this person, the officers were interviewed or had to write a report detailing the incident and their participation in it. The officers assigned to the interview and those assigned to the report conditions either conferred (the officers in the group discussed the incident, their roles, and perceptions) or did not confer before the interview/writing the report. The officers that wrote reports were to detail the incident and their role in it; the officers interviewed participated in a "cognitive interview" (I assume to assess the detail of the incident and their role in it, but no description of the interview is provided). The conflict incident was recorded by three different cameras. The interviews/reports were transcribed and three graders categorized every reported detail from the officers (e.g., including clothing noted, the person, thoughts, action) into one of Nideffer's (Nideffer & Sharpe, 1978) quadrants of attention in which the allocation of attention is defined as: external narrow focus; external broad focus, internal narrow focus, internal broad focus. Afterwards the graders viewed the video tapes and noted the officer's accuracy of the report on that item or action—(as the author states, "trying to assess the focus of attention during the *high stress component* of this scenario"). Heart rate was monitored throughout the scenario and during the interview/report writing (I assume to assess exertion, but it is not clearly stated). Also, a measure of maximum physical exertion was recorded at the end of the study. **Results.** The average number of incidents classified (by the graders) in each of the attention quadrants is reported (collapsed over the different report and confer conditions and separately for each of the report and confer conditions). Also, the average number of errors (determined by the graders) in each of the attention quadrants is reported. Finally, heart rate (HR) data is presented in a table (and analyzed) which shows the minimum, maximum and average beats-per-minute when officers were exposed to the shotgun fire compared to their HR during the interview. The HR data is the only data in which statistical comparisons were made. A significantly greater peak HR was found during the shot-gun fire in the scenario compared to when the officer was being interviewed—a correlation was also conducted on same data. No other comparisons were statistically evaluated for HR and hence no other claims related to HR can be validly or reliably drawn. This includes claims concerning HR and its affects on memory recall because no statistical comparisons were made to examine if variations in HR are related to differences in memory recall.

There are several flaws in this study that render the study invalid and unreliable. First, there are no statistical comparisons of the means across the different attentional quadrants and hence no valid or reliable claims can be made concerning mean differences across these quadrants. That is, one cannot reliably claim that officers were engaged in "a particular attentional process" more frequently than another (e.g., external focused attention vs. external broad attention). Also, due to the lack of statistical comparisons between the different reporting conditions (interview/written report) and confer

conditions (confer/no confer), one cannot draw any valid or reliable conclusions concerning any similarities or differences between any of these conditions. The author claims that statistics relevant to these means are presented in Addendum 2, but there are no statistics reported that are relevant to comparing these means and determining whether any means are statistically different from one another. Second, another flaw in the study that makes the author's claims invalid (and hence unreliable) is the lack of a control condition or control group in which attention (assessed by memory recall) is measured during a non-conflict event. Such a control condition (or group) is needed to serve as a baseline in attentional processing (i.e., memory recall of an event) in order to determine if the allocation of attention (ability to recall information) in a conflict event (stressful situation) is different than that in a non-conflict event (less stressful situation). The author states, however, that the:

"External Narrow quadrant of attention has generally nine times more data recorded in it than the Internal Broad quadrant, and the Internal Narrow quadrant is generally over twice as large as the external broad quadrant. Therefore, the researchers can **state unequivocally that during the most stressful component** of this simulation the constables [officers] experienced a narrow focus of attention and were externally tunneled specifically on items or behavior that were crucial for their performance in this situation" (page 120).

Because the number of incidences recalled or number of errors was not compared across different time points of the scenario (stressful gun shot vs. other events), one cannot validly (reliably) determine whether recall information classified under the External Narrow quadrant of attention was greater in the most stressful component of the task compared to other components of the task—or to less stressful events in general. Third, the statement above that *the officers were externally tunneled specifically on items/behavior that were crucial for their performance*, cannot be evaluated, and hence is not reliable. There is no mention in the design that officers reports were recorded into one of the four attention quadrants only if they were crucial for their performance—or that any memory recall data was coded or analyzed based on whether or not it was crucial to performance. Fourth, there is no mention or statistical analysis showing that the three "graders" assigned the memory-recall events consistently to the different attentional quadrants; that is, inter-rater reliability was not assessed, which is done to ensure reliability across the raters and hence validate the measurement used. Instead, it appears that measurements of all raters were summed together in each of the quadrants and a mean was obtained based on these summed scores—hence rater variability can confound the results and hence make any findings unreliable. Finally, it is important to mention that while accurate memory recall for events suggests that this information was previously attended, it does not mean that items which are not recalled from memory were not attended. The author erroneously makes the claim that what is recalled or not recalled (or not correctly recalled) can be used as a basis of what was and was not attended. Such a claim is not valid or reliable based on the methods used in this study because memory recall was measured which does not directly assess the allocation of attention. There are other problems concerning the sample size and unequal observations in each condition that also challenge the validity of this study. However, just having one of the violations described above is enough to show that the data and interpretations drawn from this data are unreliable.

Attachment 2

Newsletter from "Force Science News" #176 and an online link to the Discovery Channel which features "The Force Science Exhaustion Study" conducted by William Lewinski and colleagues which attempts to examine the relationship between memory recall and exhaustion.

Summary of study: In a sample size of 52, memory of 26 officers who were assigned to be exerters (exerted themselves by attacking a bag with assault movements) was compared to 26 officers who were assigned to be partners (to observe as actions took place—a non-exertion control group). All officers were told to stay alert and try to absorb and remember as much as they could. Also, all officers were given a crime report to read that described an armed robbery crew that had attacked 3 locations. Afterwards, the exerter punched a bag until exhausted while the partner (control) watched, then both the exerter and partner ran to a trailer where they were told a "known felon" was suspected of occupying. During the active scenario, different events/objects were encountered. It is claimed that heart rate and lactate levels were greater for exerters vs. partners, suggesting that exerter's bodies were working harder and were more exhausted than the partners. However, no statistics on these measures are mentioned, so the validity (reliability) of this claim is questionable. Some of the results were as follows. When estimating within 90% the number of each type of blow to the bag, 89% (or 23) of exerters estimated within the 90% accuracy rate the number of elbow, knee and palm strikes, compared to 45% (or 11) of the partners (observers). Lewinski explains, that there were very few of these types of strikes; and that with the most common hits (punches), 25% fewer exerters vs. partners were able to accurately estimate fist blows (note that 25% fewer is likely a small number of individuals because 100% of the partners were not likely accurate). Also, a statement is made by one of the researchers that "the *more* exhausted officers were, the less accurate their estimates tended to be". Importantly, **there is no data mentioned or information in terms of statistics to validate or show that this claim is reliable** (the claim suggests that a significant negative correlation was found between what was remembered and the level of lactate or heart rate). Although it is reported that exerters and partners (observers) were equally accurate in reporting threat cues (e.g., weapons) and a threatening individual, 90% of observers (or ~23) could name one descriptive item of a person in route to a target location while 1/3 of the exerters could not (suggesting that 2/3 or ~17 exerters could name one descriptive item). Finally, 27% (or 7) of exerters correctly identified the suspect (in a picture line-up) compared to 54% (or 14) of partners (observers). Again, however, no statistics are referenced to validate and ensure reliability of any of these claims that exerters performed more poorly in their memory reports than partners (the control group). Also, there is no manipulation or measurement in this study to validate the claim by Lewinski that "the ability to fully shift attention is inhibited, so even some potentially relevant information tends to get screened out". Moreover, the officers' ages varied greatly and were not matched in terms of age or baseline heart rate before assigning them to the different conditions (e.g., exerter/non-exerter). **Because the samples represented in the different conditions differed before the start of the experiment, it is not possible to draw any conclusions between these conditions.** This major confound makes this data and any conclusions drawn from it invalid and unreliable. Finally, **because this study is only assessing memory, one cannot make a statement about what was NOT attended.** It is possible that this information was attended, but was forgotten. In other words, what is explicitly remembered does indicate that this information was attended, but what is attended does not determine what will be remembered. In summary, one can not reliably draw conclusions from this study due to lack of reliability established through statistical analyses and due to some of the claims made concerning attention that cannot be validated since memory (not attention) was measured—these two cognitive processes are not identical.

Attachment 3

Evaluation of Dissertation Document by William J. Lewinski (published by UMI)

1. Overall analysis: More of an introspective, philosophical text guided toward self-promotion and development of a product for marketing (slide presentation on training techniques developed by the author) with no empirical research to validly show or even suggest that the marketing tool would be effective in improving training of police officers.
2. The only research described in this document, is based on a class study (students taking a course titled "Human Factors" at Mankato State University) in which the author was the instructor. In this course, students were given or conducted the following in this order (see pages 56-57) paraphrased here: Slide presentation on nature of course and training techniques, discussion of Zen approach to crisis training, introduction to sports psychology and the concept of peak performance and training processes by Russians and Americans which included readings from 3 different texts, practice with physiological and psychological components of peak performance (e.g., breathing techniques), participation in a ropes course, further practice with the cognitive and physiological components of peak performance, focusing on performance on the ropes course as a training goal, a second attempt at the ropes course, evaluation of the ropes course experience, extending application of course concepts and reading, participation in a mock oral board interview. This class study is the main basis for the author's claim that the marketing tool he has designed (slide presentation on training) will improve the training of police officers. The author states (see page 60):

"The evidence for the students' increased facility in the use of the training techniques was vastly apparent on the second trip on the ropes course. Every single student significantly increased the speed with which he/she traversed the course, subjectively expressed much more confidence while on the course, and demonstrated an obvious feeling of accomplishment after completing the course. Every student also noted a significant increase in his/her ability to do the skill tasks required on the course, as well as their overall enjoyment while traversing the ropes".

However, there are several major problems with this class study that jeopardize its validity.

First, when describing the results of this class study (p. 60), there is no quantitative or qualitative data reported and no statistical analyses to support the findings that every single student **significantly** increased the speed in which he traversed the ropes course the second time. More importantly, even if the data and analyses were presented and had supported the author's claims, the research design has a major flaws that compromise the validity of the study and hence any interpretation of the data. The first flaw (which is a major confound) is that there was **no control group**. Thus, one cannot determine whether the improved time on the ropes course the second time was not simply due to practice (i.e., having already done the ropes course once prior should lead to faster performance a second time due to practice compared to traversing the ropes course

the first time) or due to exposure to the training slide show and other content covered in the course. That is, practice alone could account for this improved performance and not exposure to the training slide show or contents of the class. One would need a comparison group of participants that were not in the course (and hence did not see the training slide show and were not exposed to the course content) to also perform the ropes course twice. Only if it were then found that the students participating in the class showed a greater improvement in time to complete the ropes course (time to complete the ropes course the second time minus the time to complete the ropes course the first time) compared to the participants who did not take the class (control group) could one attribute the improvement to the course content (as opposed to practice on the course). **This is the most serious confound a study can have that basically invalidates any interpretation of the data.** It is referred to as a “testing confound” and refers to repeated measurement of an individual which can lead to learning or strategy transfer as well as other effects. Such a confound invalidates any interpretation of data.

There are other major flaws (biases) in this study as well that challenge its validity. First, it is likely that the subjective reports from the students were influenced by demand characteristics. That is, responses by the participants may be influenced by how they think they should feel or act. They are being evaluated by their instructor, and hence may want to please the instructor or do well in the class, and this could influence their responding. Also, they are reading and being exposed to information that suggests they should act or feel a particular way after engaging in certain activities, hence they may falsely report that this is what they felt to be part of the norm. One way to help prevent demand characteristics is to not let the participants know what might be the expected outcome. Second, **experimenter bias** could likely have influenced the outcome of the data or interpretation of the data observed by the experimenter. This can occur either consciously or unconsciously on the part of the experimenter in order to support his hypothesis or theory. One way to combat this possible bias is to have someone collect and analyze the data who is not aware of the hypothesis being tested. The fact that no data or data analyses are presented with conclusions simply based on “I observed” makes it very difficult to rule out experimenter bias.

Separate from the major flaws described above, there is also another major flaw in this study that challenges its validity. The claim by the author is that the slide show on training techniques alone improved student performance/confidence. The problem here is that the classroom study reported exposed students to the training slide show as well as to other information, reading resources, and training experiences. Thus, it is difficult to assess whether the training slideshow alone, the slideshow together with the other experiences, or one or more of the other experiences alone led to these conclusions. Because other factors besides the slideshow could have affected performance, there is no way to know how the slideshow specifically influenced performance (even if all other confounds in this study had been controlled for). In short, there is no evidence that the

training slideshow should improve performance, and hence there is no evidence that this training slideshow may improve officer training (which is the argument the author is making in this text).

Finally, **the premise of the author's manuscript** is that "police officers are not well trained". However, there is no empirical evidence presented by the author that this is the case. Also, there is no evidence that the training slideshow to be marketed can improve performance in a stress-full task (i.e., on a ropes course). This study assessing the effectiveness of the training slideshow has many major confounds that make it invalid. In fact, this study leads to false empirical claims that this training slideshow can improve officer training.